

HOLLAND & KNIGHT LLP

**701 Brickell Avenue
Suite 3000
Miami, Florida 33131
(305) 789-7773**

**Application
for
United States
Letters Patent**

filed on behalf of

**Applicant(s): Vadakkedathu T. Rajan
Jason L. Crawford
Mark N. Wegman**

**For: Classification of Electronic Mail into
Multiple Directories Based upon
Their Spam-Like Properties
Attorney Docket: YOR920030461US1**

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

**CLASSIFICATION OF ELECTRONIC MAIL INTO MULTIPLE
DIRECTORIES BASED UPON THEIR SPAM-LIKE PROPERTIES**

BACKGROUND OF THE INVENTION

Field of the Invention.

[0001] The invention relates to processing unsolicited electronic mail. More particularly, this invention relates to classifying electronic mail into multiple directories based upon their spam-like properties.

Description of the Background Art.

[0002] Electronic mail ("e-mail") is a common form of communication for many individuals and most organizations throughout the world. Unfortunately however, e-mail recipients are often the targets of unsolicited mass mailings sent by mass marketers offering their products or services. Reminiscent of excessive mass solicitations via postal services, facsimile transmissions and telephone calls, an e-mail recipient may receive hundreds of unsolicited e-mails over a short period of time. Similar to the task of handling "junk" postal mail and faxes, an e-mail recipient must laboriously sift through his or her incoming mail simply to

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

sort out the unsolicited "junk" e-mail from legitimate e-mails. Such unsolicited "junk" e-mail is in common parlance referred to as "spam".

[0003] Many vendors of electronic mail servers, as well as many third-party vendors, offer spam-blocking software to detect, label and sometimes automatically remove spam. Presently, there exist many methods for detecting, labeling and removing spam. Representative methods are taught in the following U.S. patents, the disclosures of which are hereby incorporated by reference herein:

Patent No.	Title
5,999,932	System and Method for Filtering Unsolicited Electronic Mail Messages Using Data Matching and Heuristic Processing
6,023,723	Method and System for Filtering Unwanted Junk E-Mail Utilizing a Plurality of Filtering Mechanisms
6,029,164	Method and Apparatus for Organizing and Accessing Electronic Mail Messages Using Labels and full Text and Label Indexing
6,092,101	Method for Filtering Mail Messages for a Plurality of Client Computers Connected to a Mail Service System
6,161,130	Technique Which Utilizes a Probabilistic Classifier to Detect "Junk" E-Mail by Automatically Updating A Training and Re-Training the Classifier Based on the Updated Training List
6,167,434	Computer Code for Removing Junk E-Mail Messages
6,199,102	Method and System for Filtering Electronic Messages
6,249,805	Method and System for Filtering Unauthorized Electronic

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

	Mail Messages
6,266,692	Method for Blocking All Unwanted E-Mail (Spam) Using a Header-Based Password
6,324,569	Self-Removing Email Verified or Designated as Such by a Message Distributor for the Convenience of a Recipient
6,330,590	Preventing Delivery of Unwanted Bulk E-Mail
6,421,709	E-Mail Filter and Method Thereof
6,484,197	Filtering Incoming E-Mail
6,487,586	Self-Removing Email Verified or Designated as Such by a Message Distributor for the Convenience of a Recipient
6,493,007	Method and Device for Removing Junk E-Mail Messages
6,654,787	Method and Apparatus for Filtering E-Mail

[0004] As taught by several of the above-referenced patents, there exist spam filters that grade the spaminess of incoming e-mail by processing the e-mail for spam-like properties along a scale (e.g. 0-100) and if the incoming e-mail is graded to have a spaminess level above a predetermined numeric threshold (e.g., above 80), the e-mail is automatically moved from the recipient's Inbox into a spam directory. Ideally, all of the spam will be moved to the spam directory, thereby obviating the need for the recipient to read the e-mail in the spam directory. As used herein, the term "spaminess" may include undesired or unsolicited e-mails determined on a variety of objective and subjective scales including but not limited to politics, pornography and marketing

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

scams.

[0005] Many algorithms exist for processing incoming e-mail and grading the spaminess of the e-mail. Some representative algorithms are taught in the foregoing patents whereas others are taught in the following publications, the disclosures of which are hereby incorporated by reference herein.

Hooman Katirai, Filtering Junk E-Mail: A Performance Comparison between Genetic Programming and Naïve Bayes, September 10, 1999.
Jefferson Provost, Naïve-Bayes vs. Rule-Learning in Classification of Email, Technical Report AI-TR-99-284.
Mehran Sahami, Susan Dumais, David Heckerman & Eric Horvitz, A Bayesian Approach to Filtering Junk E-Mail, www.paulgram.com
Paul Graham, Stopping Spam, August 2003, www.paulgram.com
Paul Graham, So Far So Good, August 2003, www.paulgram.com
Paul Graham, Filters That Fight Back, August 2003, www.paulgram.com
Paul Graham, Better Bayesian Filtering, January, 2003, www.paulgram.com

[0006] Unfortunately, it is often the case that the predetermined threshold, coupled with the inherent inaccuracies of the algorithm employed for grading of the incoming e-mail for spaminess, results in some e-mail being misclassified as spam when it is not, or visa versa. Obviously, as incoming mail is assigned a high score based upon its level of

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

spaminess, the likelihood of classifying incoming e-mail as spam increases as the threshold is decreased. However, this disadvantageously results in a greater likelihood of non-spam e-mails being misclassified as spam and consequently being overlooked and not read by the recipient. Conversely, increasing the threshold decreases the chance that non-spam e-mails are misclassified as spam.

[0007] The use of a spaminess threshold thus results in a paradox of being, on the one hand, too guarded of potentially misclassifying non-spam e-mails as spam by raising the threshold too high whereupon a significant number of spam e-mails would fail to be identified as spam and remain in the recipient's Inbox and, on the other hand, being too aggressive by reducing the threshold resulting in non-spam e-mails being classified as spam. Consequently, the recipient is often faced with the dilemma of having an inbox with significant amounts of spam or having to frequently scan the presumed spam e-mail in the spam directory to verify that a legitimate e-mail was not improperly moved to the spam directory.

[0008] Therefore, it is an object of this invention to provide an improvement which overcomes the aforementioned inadequacies of the

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

prior art and provides an improvement which is a significant contribution to the advancement of the art of filtering spam.

[0009] Another object of this invention is to provide an apparatus and method for handling incoming e-mail based upon the e-mail's spam-like properties.

[0010] Another object of this invention is to provide an apparatus and method for grading incoming e-mail for spaminess and, based upon the e-mail's spaminess grade, moving the incoming e-mail to at least one of a plurality of appropriately labeled directories.

[0011] The foregoing has outlined some of the pertinent objects of the invention. These objects should be construed to be merely illustrative of some of the more prominent features and applications of the intended invention. Many other beneficial results can be attained by applying the disclosed invention in a different manner or modifying the invention within the scope of the disclosure. Accordingly, other objects and a fuller understanding of the invention may be had by referring to the summary of the invention and the detailed description of the preferred embodiment in

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

addition to the scope of the invention defined by the claims taken in
conjunction with the accompanying drawings.

SUMMARY OF THE INVENTION

[0012] For the purpose of summarizing this invention, this invention comprises an apparatus and method for providing, in addition to the conventional "Inbox" directory in which all incoming mail is normally received, a plurality of appropriately labeled directories for containing e-mails suspected of being spam, grading the level of spaminess of the incoming e-mail and then moving or copying incoming e-mail into one or more of the spam directories based upon the e-mail's respective level of spaminess.

[0013] For example, according to the invention, each piece of incoming mail is graded along a scale, according to any of the prior art methods, to determine the level of spaminess of the e-mail. At least two spam directories are provided and labeled to contain spam of certain ranges of spaminess along the scale. Each piece of incoming mail is graded (objectively by an algorithm and/or subjectively by "voting and teaching" by the recipient, or a combination thereof) and, if determined to possibly or likely constitute spam, then moved into the respective directory according

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

to its level of spaminess. Upon grading, if the incoming e-mail is determined with a high degree of probability not to constitute spam, it is left in the Inbox.

[0014] For example, a spam directory labeled "black" could be provided for containing e-mail whose level of spaminess is graded high, representative of likely constituting spam. If the scale employed is 0 to 100, an exemplary range could be above 80 for the "black" spam directory. Correspondingly, a spam directory labeled "grey" could be provided for containing e-mail whose level of spaminess is graded medium, representative of probably constituting spam. An exemplary range for the "grey" spam directory could be 30 – 80. Thus, all incoming mail graded to have a spaminess level of above 80 would be moved from the Inbox to the "black" spam directory, all mail graded to have a spaminess level of 30-80 would be moved to the "grey" spam directory and, finally, all mail graded to have a spaminess level less than 30 would remain in the Inbox.

[0015] The invention advantageously grades e-mail according to the level of spaminess and then moves the e-mail into the appropriately labeled directory (as used herein, a directory may optionally be virtual

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

directory or may reside in a separate repository in the same server or in a remote server). In this manner, the recipient has a high degree of confidence that all of the mail in his or her Inbox will not constitute spam. More importantly, the recipient may review the mail suspected to constitute spam in each of the spam directories at a frequency and with a level of care commensurate with the assigned spaminess range associated with each spam directory. Indeed, the recipient may not necessarily review the mail in the directory with the highest level of spaminess. Moreover, such mail may be automatically deleted immediately or with an appropriate delay.

[0016] More specifically, when employing the above example, the recipient's Inbox may be regarded as the "white" directory containing e-mail that determined with a high degree of certainty to not constitute spam. The recipient may constantly monitor the "white" Inbox for incoming e-mail and read the mail being received therein without being distracted by the existence of spam. The recipient may choose to read the e-mail in "grey" directory less frequently, such as only at the close of each business day, to determine if any of the "grey" mail therein is legitimate

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

mail and needs to be responded to or otherwise handled by the recipient. Other algorithms may be employed to allow the recipient to vote on the grey mail and train the system to better determine spaminess according to the recipient's personal preferences. Finally, the recipient may choose to infrequently (e.g., once a week) but quickly scan through the headers of the e-mail in the "black" directory to verify that there is no legitimate e-mail therein.

[0017] It should be appreciated that greater granularity beyond the "white" Inbox, the "grey" spam directory and the "black" spam directory can be achieved according to the subject invention by simply including a greater number of spam directories, each assigned with their respective range of spaminess and with each being appropriately labeled to indicate to the recipient how often and with what degree of care the spam e-mail therein should be read or scanned to verify that no legitimate e-mails are inadvertently overlooked. Moreover, the labeling of the Inbox and spam directories could be done with color-codes. For example, in the case of a more graduated five ranges, the colors could be Green (for the Inbox) and then Blue, Yellow, Orange and Red (for the first, second, third and forth

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

spam directories) with five assigned a spaminess ranges of 0-20, 21-40, 41-60, 61-80 and 81-100, respectively.

[0018] It should also be appreciated that without departing from the spirit and scope of the invention, the spaminess ranges assigned to the adjacent directories may overlap and each e-mail whose grade is included within the overlap, is also copied into the adjacent directory. For example, using the foregoing example, the five graduated ranges for the Green, Blue, Yellow, Orange and Red directories, may be assigned spaminess ranges of 0-20, 21-40, 41-60, 55-85 and 81-100, respectively such that the Orange span directory overlaps with the Yellow and the Red spam directories. Thus, an e-mail having been graded an 85 would be moved from the Inbox into the Red spam directory but also copied into the Orange spam directory where it is likely to be given more scrutiny by the recipient that it otherwise would be given if it was only included in the Red spam directory. Similarly, an e-mail graded with a grade of 58 would be included in both the Yellow and Orange spam directories.

[0019] The foregoing has outlined rather broadly the more pertinent and important features of the present invention in order that the detailed

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

description of the invention that follows may be better understood so that the present contribution to the art can be more fully appreciated. Additional features of the invention will be described hereinafter which form the subject of the claims of the invention. It should be appreciated by those skilled in the art that the conception and the specific embodiment disclosed may be readily utilized as a basis for modifying or designing other apparatuses and methods for carrying out the same purposes of the present invention. It should also be realized by those skilled in the art that such equivalent apparatuses and methods do not depart from the spirit and scope of the invention as set forth in the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] For a fuller understanding of the nature and objects of the invention, reference should be had to the following detailed description taken in connection with the accompanying drawings in which:

[0021] Fig. 1 is a block diagram of the apparatus of the invention showing an e-mail Inbox and a plurality of spam directories, each containing e-mails with respective level of spaminess; and

[0022] Fig. 2 is a block diagram of the method of the invention showing a process flow of an incoming e-mail.

[0023] Fig. 3 is a screen shot showing a plurality of email folders according to an embodiment of the invention.

[0024] Similar reference characters refer to similar parts throughout the several views of the drawings.

DETAILED DESCRIPTION

[0025] Referring to Fig. 1, electronic mail ("e-mail") sent between or among individuals and companies throughout the world travels via the Internet, generally indicated by numeral 10, composed of a plurality of interconnected backbone and other sites that transfer mail from one mail server 12a uniquely identified by an Internet protocol ("IP") address to another mail server 12b uniquely identified by another IP address, each IP address being conveniently mapped to an upper-level domain name. Each domain mail server 12 typically includes one or more client mailboxes 14.

[0026] Each client mailbox 14 includes at least an Inbox 16 into which is delivered incoming e-mail addressed to the owner of the Inbox 16 as the intended recipient of the e-mail (hereinafter "recipient"). The recipient reads the e-mails contained within the Inbox 16 and then typically deletes the e-mail or copies or moves it to other directories within the mailbox.

[0027] According to the present invention, in addition to the Inbox, at least two spam directories (SPAM1, SPAM2, ...SPAM_n, $n \geq 2$) are provided. Each spam directory is assigned a spaminess range (RANGE1,

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

RANGE2,...RANGEn) within a predefined spaminess scale and appropriately labeled (LABEL1, LABEL2,... LABELn) corresponding to the increasing respective spaminess ranges RANGEn. Without departing from the spirit and scope of the invention, one or more of the ranges RANGEn may overlap with respective adjacent ranges.

[0028] As shown in Fig. 2, as each incoming e-mail (MAIL1, MAIL2,...MAILn) is received by the mail server 12 at block 20, it is graded (MAIL1GRADE1, MAIL2GRADE2, ... MAILnGRADEn,) for spaminess at block 22 and, at block 24, left in and/or moved and/or copied to the Inbox 16 and/or the spam director(ies) SPAMn to which is assigned the respective RANGEn encompassing the spaminess grade MAILnGRADEn.

[0029] It shall be understood that the claimed invention is operable with any algorithm that grades e-mails for spaminess. The algorithms employed for grading the incoming mail MAILn may therefore constitute any prior art algorithm, or any future developed algorithm, that grades e-mails for spaminess. It shall also be understood that the mail system employed may comprise a client/server topology in which event the computing involved for grading the e-mail MAILn for spaminess may be

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

processed either on the server or the client processors, or partly by both.

Furthermore, in the event of processing on a mail server or other computing topology, the incoming mail MAIL_n once graded may bypass the Inbox altogether and be immediately moved or copied into the respective spam directory SPAM_n according to its grade MAIL_nGRADE_n. Finally, it shall be understood that during set-up the recipient or his or her stead may specify the number of spam directories desired. Additional user-settable configurations may include the ability to name and color-code the spam directories, as well as the ability to assign their respective ranges thereto, with or without any overlapping.

[0030] Fig. 3, shows a screen shot showing a plurality of email directories (or folder) according to an embodiment of the invention. Under the 3MGFilter directory are found the grey, black, voted spam, and pre-inbox folders. According to the embodiment illustrated in Fig. 3, email whose spaminess (content of spam-like characteristics) has been determined by the mail server to be so high that the probability of it being non-spam (white mail) is almost zero is placed in the black folder. The directory labeled gray contains email that is strongly suspected of being

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

spam but there is a definite chance that it is not spam. The user may choose to delete spam in the black directory with little or no inspection. However, it is preferred that the user inspect the email in the gray folder before deleting any of the mail therein. Moreover some email may be placed in more than one directory such as the inbox (white) directory and the gray directory. The email placed in more than one directory is that which contain characteristics that place it in two categories or in an overlapping range. Preferably, the email place in more than one directory would include a visual indication to the user that the email is contained in more than one directory and possibly prompting the user to take some action such a voting on its spaminess. The system is preferably a smart system that uses the user's votes throughout the network as a factor in determining the spaminess scores of other email. Thus, the system is adaptive to new information acquired from handling email in the directories.

[0031] According to another embodiment, as the system learns more about the received email, it can move emails from one directory to another. For example, if a user votes an email as spam all email resembling the

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

email that was the subject of the vote is moved from the gray directory to the black. Alternatively where the classification of an email is not certain in light of recent information the system could move it and other similar emails from the inbox to the gray directory.

[0032] The system can base its decision on which directory receives an email on criteria other than spaminess. For example, the system can mine a personal profile of the user to determine which is the appropriate directory for an email. Thus, all email from the user's spouse would be placed in the appropriate folder or all email relating to a pastime of the user such as tennis could be placed in a non-spam directory. Changes in the user profile would affect the decision as well. Other criteria for directory selection include the degree of offensiveness. The system can create multiple directories based on user behavior.

[0033] The decision of where to store received emails is not just made when they received. As the system modifies its criteria based on new information observed from the user's behavior or other data it can re-classify email already received and possibly relocate the emails to the newly-appropriate directory or otherwise prompt the user to reconsider

PATENT

EXPRESS MAIL NO. EL968293664US

Docket No. YOR920030461US1

items of received email. Hence, new variables can be introduced to the classification algorithm as new considerations as determined to be relevant. Additional classification criteria may be obtained by allowing the user to adjust the level of spaminess that is required for placement of emails in say the gray or black folders. That allows for more detailed control for a user who wants to have a say in the risk involved in deleting spammy email without looking.

[0034] The present disclosure includes that contained in the appended claims, as well as that of the foregoing description. Although this invention has been described in its preferred form with a certain degree of particularity, it is understood that the present disclosure of the preferred form has been made only by way of example and that numerous changes in the details of construction and steps and the combination and arrangement of parts and steps may be resorted to without departing from the spirit and scope of the invention.

[0035] Now that the invention has been described, we claim: